

Relational Entropic Dynamics of Particles*

Selman Ipek and Ariel Caticha

Department of Physics, University at Albany–SUNY, Albany, NY 12222, USA

Abstract

The general framework of entropic dynamics is used to formulate a relational quantum dynamics. The main new idea is to use tools of information geometry to develop an entropic measure of the mismatch between successive configurations of a system. This leads to an entropic version of the classical best matching technique developed by J. Barbour and collaborators. The procedure is illustrated in the simple case of a system of N particles with global translational symmetry. The generalization to other symmetries whether global (rotational invariance) or local (gauge invariance) is straightforward. The entropic best matching allows a quantum implementation Mach’s principles of spatial and temporal relationalism and provides the foundation for a method of handling gauge theories in an informational framework.

1 Introduction

The question of whether motion is absolute or relative has been around since the beginning of mechanics. At first the issue was how to define velocity in order to define motion, but ultimately the issue is how to define change in general. Newton’s solution was to describe motion in terms of the evolving coordinates of particles embedded in an absolute space and in an absolute time but objections were immediately raised because, as Newton himself realized, neither absolute space nor absolute time are observable. A presumably “better” mechanics would describe motion not in terms of changes of unobservable absolute positions but in terms of the observable *relative* distances between them. The search for such *relational* forms of mechanics eventually led to Einstein’s general relativity and for a while the nature of relative motion was thought to be fully understood — at least within the context of classical physics.

Quantum mechanics, however, raises new questions. One problem is that the quantum version of our best relational theory — general relativity — does not yet exist. Another is that, in its standard formulation, quantum mechanics is manifestly non-relational: it lives in Newton’s absolute space and time, or at

*Presented at MaxEnt 2015, the 35th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (July 19–24, 2015, Potsdam NY, USA).

best in Minkowski's absolute space-time. Furthermore, in the standard Copenhagen interpretation quantum particles do not have definite positions much less definite relative distances; no such theory could ever be relational in the usual sense.

Our goal is to take the first steps towards a relational quantum mechanics by using some of the recent conceptual innovations introduced by Entropic Dynamics (ED) [1]-[4]. ED is a framework for formulating dynamical models as applications of well established principles of inference [5] and, as one might expect, the tools for inference — probabilities, entropies, and also information geometry — play a dominant role. It is indeed a very appealing aspect of ED that its very foundation emphasizes the close relationship between quantum theory, geometry, and inference. This suggests that ED might provide the proper setting for the unification of quantum theory and gravity. Some preliminary steps in this direction are already being taken in [6] and [7].

As with all applications of entropic methods ED requires that we first specify the subject that is the goal of our inferences — the microstates. Then we must identify the relevant information on the basis of which the inference is to be carried out — the constraints. Thus, the first step towards a relational ED is to specify relational microstates and to this end we make extensive use of the insights into classical versions of relational mechanics achieved in the pioneering work of Julian Barbour and his collaborators [8]-[9].

The basic problem is that one cannot go very far in formulating a relational mechanics in terms of the relative interparticle distances — it is just not practical. Thus, on one hand one is forced to rely on particle coordinates, and on the other hand, the coordinate description fails because it is redundant. Indeed, two configurations that differ by an arbitrary displacement or an arbitrary rotation describe exactly the same physical situation.

To handle the redundancy Barbour and Bertotti invented a technique they called best matching (BM) [8]. The idea is to introduce some quantitative measure of the mismatch or the “distance” between two successive configurations and then shift and rotate them to find the position of one configuration *relative* to the other that minimizes the mismatch. Then the actual intrinsic change is defined through the least mismatch between successive configurations as they are subjected to translations and rotations relative to each other.

The Barbour-Bertotti best matching procedure is crucial to implement spatial relationalism (Mach's first principle) and also temporal relationalism (Mach's second principle). The former abolishes absolute space. The latter, which abolishes absolute time, is the notion that physical changes do not occur *in an external absolute time*, but rather that time is an abstraction at which we arrive from studying the changes *in things*.

The choice of mismatch measure is central to the relational program. The particular choice adopted in [8] for a relational classical mechanics is borrowed from Jacobi's action principle and amounts to a variation on a least-squares mismatch. It succeeds and it is elegant but suffers from the same flaws that one can attribute to other classical theories: Where do these action principles come from? What is distance? Why geometry? How does one justify the ad hoc

quantization rules that lead to quantum mechanics? How does one implement best matching in a quantum context?

The ED framework provides insight into all these questions. ED leads us directly to a quantum theory; no questionable preliminary detours through classical dynamics are needed. In ED, dynamics is not derived from an underlying — and therefore unjustified — action principle; instead the action principle is derived. Furthermore, since the quantum states are probability distributions we already have a unique measure of mismatch. It is given by the relative entropy and the resulting mismatch is exactly the distance given by information geometry. This leads us to a criterion of *entropic best matching* (EBM) which, as we shall see, is all we need to derive a relational ED.

We conclude this introduction with the observation that the relational program — the implementation of Mach’s principles through best matching — can be applied to more than just the relational motion of particles in space, as Mach had originally intended. An important insight by Barbour and Bertotti is that it applies to any model with redundancy in description. Such models include all fundamental theories such as electromagnetism [8], Yang-Mills theories [11], and gravity [9].

In this paper we develop the basic formalism for a relational ED and we demonstrate entropic best matching in the context of a simple model. We consider a system of N particles with the simplest redundancy — a global translational symmetry. The generalization to include global rotations is straightforward and will be treated elsewhere.

2 Entropic Dynamics

The microstates We deal with a system of N particles living in a three dimensional flat space \mathbf{X} with metric δ_{ab} . We describe the microstate of the N particle system by a point $x \in \mathbf{X}_N$ in a $3N$ dimensional configuration space, $\mathbf{X}_N = \mathbf{X} \otimes \cdots \otimes \mathbf{X}$, with coordinates x_n^a where $a = 1, 2, 3$ is the spatial index and $n = 1 \dots N$ labels the particle. (Throughout we use notation consistent with that of [3]). The specification of the microstate in terms of the coordinates x_n^a is redundant in the sense that shifting all particles by the same amount does not lead to a different physical state. In other words, x_n^a and $x_n^a + \xi^a$ where ξ^a is a constant independent of x and of n describe the same physical situation.

In ED particles are assumed to have well defined positions and the goal is to infer what those unknown positions might be; we want to assign a probability distribution $\rho(x)$. The main assumption is that the motion is continuous which means that it can be analyzed as a sequence of short steps from x_n^a to $x_n^a + \Delta x_n^a$. The method of maximum entropy is used to find the probability $P(x'|x)$ that the system will take a short step from x_n^a to $x_n'^a = x_n^a + \Delta x_n^a$. Then these short steps will be iterated to find the evolving $\rho(x, t)$.

However, in a relational dynamics x_n^a and $x_n^a + \xi^a$ represent the same initial state just as $x_n'^a$ and $x_n'^a + \xi'^a$ represent the same final state. Then the short

step is represented by

$$\hat{\Delta}x_n^a = (x_n'^a + \xi'^a) - (x_n^a + \xi^a) = \Delta x_n^a + \Delta \xi^a, \quad (1)$$

where the spatial shift $\Delta \xi^a$ is arbitrary. The two configurations are said to be “best matched” when $\Delta \xi^a$ is chosen to minimize a certain entropic measure of mismatch to be defined later. Finding the optimal $\Delta \xi^a$ amounts to establishing a criterion of “equilocality” between successive instants; it amounts to deciding which position x'^a at the later instant *is that same as* a position x^a at the earlier instant. Once equilocality has been established two successive configurations are intrinsically identical when $\hat{\Delta}x_n^a = 0$.

At this point in our argument the optimal $\Delta \xi^a$ is still unknown. However, to make progress we will assume that a notion of equilocality has been established through a trial shift $\Delta \xi^a$ to be determined later when the entropic measure of mismatch is defined.

Maximum Entropy To find the transition probability we maximize the entropy of $P(x'|x)$ relative to a prior $Q(x'|x)$,

$$S[P, Q] = - \int dx' P(x'|x) \log \frac{P(x'|x)}{Q(x'|x)}, \quad (2)$$

subject to the appropriate constraints. To represent an initial state of extreme ignorance we adopt a uniform prior.¹ Thus $Q(x'|x) \approx Q$ is a constant which can be dropped because it has no effect on the maximization.

The information about the motion is introduced through constraints. The fact that particles move by taking infinitesimally short steps is imposed through N independent constraints,

$$\left\langle \delta_{ab} \hat{\Delta}x_n^a \hat{\Delta}x_n^b \right\rangle = \int dx' P(x'|x) \left(\hat{\Delta}x_n^a \hat{\Delta}x_n^b \delta_{ab} \right) = \kappa_n. \quad (3)$$

To ensure the continuity of the motion we shall eventually take the limit $\kappa_n \rightarrow 0$. Correlations and entanglement among the particles are imposed through one additional constraint,

$$\sum_n \left\langle \hat{\Delta}x_n^a \right\rangle \frac{\partial \phi(x)}{\partial x_n^a} = \int dx' P(x'|x) \sum_n \hat{\Delta}x_n^a \frac{\partial \phi(x)}{\partial x_n^a} = \kappa', \quad (4)$$

where ϕ is a “drift” potential² which plays a role somewhat analogous to the pilot wave in de Broglie-Bohm theory and κ' is another small but for now unspecified position-independent constant.

¹Improper, non-normalizable priors are known to lead to problems. By “uniform” in a physics context we mean a distribution $Q(x'|x)$ that is sufficiently broad that its precise functional form does not affect our inferences. In curved spaces a uniform distribution is such that it assigns equal probabilities to equal volumes. In such cases $Q(x'|x) = (\det g_{AB})^{1/2} f(x'|x)$ where g_{AB} is the metric tensor and f is a sufficiently broad scalar function.

²Elsewhere, in the context of particles with spin, we see that the potential $\phi(x)$ can be given a natural geometric interpretation as an angular variable. Its integral over any closed loop is $\oint d\phi = 2\pi n$ where n is an integer.

Maximizing the entropy (2) subject to the constraints (3), (4), and normalization leads, after some manipulation, to a Gaussian process,

$$P(x'|x) = \frac{1}{Z} \exp - \sum_n \frac{\alpha_n}{2} \delta_{ab} \left(\hat{\Delta} x_n^a - \frac{\alpha'}{\alpha_n} \frac{\partial \phi}{\partial x_n^a} \right) \left(\hat{\Delta} x_n^b - \frac{\alpha'}{\alpha_n} \frac{\partial \phi}{\partial x_n^b} \right). \quad (5)$$

where Z is a normalization constant and $\{\alpha_n, \alpha'\}$ are Lagrange multipliers. Continuity is achieved imposing that $\alpha_n \rightarrow \infty$. As discussed in [10] the multiplier α' can be absorbed into ϕ , which amounts to setting $\alpha' = 1$ without changing the dynamics.

Entropic Time In a relational approach time is defined as an ordered succession of instants but these are not envisaged as being embedded in an externally given absolute time or space-time. Instead an instant is defined by an “instantaneous configuration”. In ED, which has from the very start been designed to be temporally relational an instant *is defined by* a probability distribution. In fact, an instant is a probability distribution [1].

The foundation of any notion of time is dynamics and here the dynamics is given by the transition probability $P(x'|x)$ in (5). Thus, if the distribution $\rho(x, t)$ refers to one instant t , then the distribution

$$\rho(x', t) = \int dx P(x'|x) \rho(x, t) \quad (6)$$

generated by $P(x'|x)$ defines what we mean by the “next” instant. In ED time is *constructed* instant by instant so that, given the present, the future is independent of the past.³ The construction leads to instants that are ordered and, because the transition probability $P(x'|x)$ is determined by *maximizing* an entropy, there is a natural arrow of time. We emphasize that this entropic time is already fully relational: time is the sequence of instants, time is the sequence of probability distributions.

To complete the construction of entropic time we address what is perhaps the least fundamental aspect of time: we specify the scale of time t . This amounts to specifying the interval Δt between successive instants. The criterion is convenience: time is defined so as to simplify the description of motion. For short steps the motion is dominated by fluctuations the scale of which is given by the multipliers α_n . This suggests setting

$$\alpha_n = \frac{m_n}{\hbar \Delta t}, \quad (7)$$

where m_n are particle specific constants that are eventually identified as masses, and \hbar is an overall constant that fixes the units of time relative to those of length and mass.

³We can see that by construction the dynamics is Markovian. But this is not the usual Markovian process that occurs in a pre-existing time; it is an entropic process that generates its own Markovian time as it unfolds.

In a relational classical mechanics it is the free particles that provide the prototype of a clock: particles move equal distances in equal times. In ED it is the fluctuations that provide the clock that sets the measure of time: particles undergo equal fluctuations in equal times. With this definition our Gaussian process eq.(5) becomes a Wiener process.

Short steps A generic displacement can be split into the expected drift plus a fluctuation,

$$\Delta x^A = \hat{\Delta} x^A - \Delta \xi^A = \langle \Delta x^A \rangle + \Delta w^A, \quad (8)$$

where the capital index $A = (n, a)$ includes both the particle n and the spatial index a so that $x_n^a = x^A$. To simplify the notation we write the spatial shift as $\xi^A = \xi^a$ noting that ξ^A is independent of the particle label n . We also introduce the mass tensor,

$$m_{AB} = m_n \delta_{AB} \quad \text{and its inverse} \quad m^{AB} = \frac{1}{m_n} \delta_{AB}, \quad (9)$$

which, as shown [3][4], is the metric of configuration space up to an unimportant scale factor.

From (1) and (5) we find that the shift $\Delta \xi^A$ affects the expected steps,

$$\langle \Delta x^A \rangle = \langle \hat{\Delta} x^A \rangle - \langle \Delta \xi^A \rangle = \hbar \Delta t m^{AB} \partial_B \phi - \Delta \xi^A, \quad (10)$$

but does not affect the fluctuations,

$$\hat{\Delta} w^A = \Delta w^A \quad \text{with} \quad \langle \Delta w^A \rangle = 0 \quad \text{and} \quad \langle \Delta w^A \Delta w^B \rangle = \hbar \Delta t m^{AB}, \quad (11)$$

which remain large $\Delta w \sim O(\Delta t^{1/2})$ and essentially isotropic. This leads us to expect, and we shall later confirm, that the optimal shift $\Delta \xi$ is of order Δt so that $\Delta \xi \ll \Delta w$.

Fokker-Planck Equation The dynamical equation of evolution, eq.(6) can be rewritten as a Fokker-Planck (FP) or continuity equation [5],

$$\partial_t \rho(x, t) = -\partial_A [\rho(x, t) V^A(x, t)], \quad (12)$$

where V^A is the velocity of the probability flow, or current velocity,

$$V^A(x, t) = m^{AB} \partial_B \Phi(x, t) - \dot{\xi}^A \quad \text{where} \quad \Phi = \hbar(\phi - \log \rho^{1/2}), \quad (13)$$

and $\dot{\xi}^A = \Delta \xi^A / \Delta t$.

The current velocity receives three types of contributions. The first two are the familiar drift and osmotic velocities described through the gradient of the “phase” Φ [4]. The third contribution is the shift velocity $\dot{\xi}^A$; it is the term that implements relationality.

3 Entropic Best Matching

We are now ready to introduce entropic best matching (EBM). The basic idea is that once we have decided on the relevant information necessary for predicting future behavior we can say that this information defines what we mean by an “instant”. In ED the relevant information is given by the distribution $\rho(x, t)$ and the (suitably updated) drift potential $\phi(x, t)$ which determines the transition probability $P(x', t'|x, t)$. An alternative representation of exactly the same information is given by the *joint* distribution⁴

$$\rho(x', t'; x, t) = P(x', t'|x, t)\rho(x, t) . \quad (14)$$

This representation is more convenient for two reasons. First, the joint distribution (14) is more suggestive of the flow of time from t to t' . Indeed, while the pair of functions $\rho(x, t)$ and $\phi(x, t)$ are assigned to a single sharp instant t , the single function $\rho(x', t'; x, t)$ refers to two instants t and t' . In fact, $\rho(x', t'; x, t)$ describes evolution from the initial instant t not just into one other instant t' but into all instants in the immediate future of t . This is precisely what we need to explore the dynamical effect of any trial shift $\Delta\xi^A$.

Second, and equally significant, is the fact that $\rho(x', t'; x, t)$ is a probability distribution while $\phi(x, t)$ is not. This means that there exists a unique criterion for quantifying the mismatch between any two such distributions; it is given by their relative entropy. For distributions that differ only slightly, as we expect in the case of *successive* instants, this is equivalent to the information distance between them.⁵ Therefore the mismatch that is relevant to our discussion is the information distance between $\rho(x', t'; x, t)$ and $\rho(x', t' + dt; x, t)$.

To summarize: A natural implementation of EBM within a framework of inference requires that we first identify which distributions are to be compared; then, using their relative entropy, we quantify their mismatch as an information distance. The *intrinsic* mismatch is the minimum distance between the distributions as they are shifted relative to each other by varying the shift $\Delta\xi^A$. The optimal shift $\Delta\xi^A$ implements equilocality.

Information Geometry and Temporal Distance The information distance dT between $\rho(x', t'; x, t)$ and $\rho(x', t' + dt; x, t)$ is given by $dT^2 = Gdt^2$ where the metric tensor G has a single component,

$$G = C \int dx dx' \rho(x', t'|x, t) [\partial_{t'} \log \rho(x', t'|x, t)]^2 , \quad (15)$$

where C is an arbitrary overall constant. Using (14),

$$G = C \int dx \rho(x, t) \int dx' P(x', t'|x, t) [\partial_{t'} \log P(x', t'|x, t)]^2 , \quad (16)$$

⁴This is not the probability of t and t' ; perhaps a better notation would be $\rho(x', x|t', t)$ where t and t' are parameters.

⁵For reviews on information geometry including the proof of uniqueness of the information metric, see *e.g.*, [5][12].

and then (5) gives

$$G = \frac{3NC}{2(\Delta t)^2} + \frac{C}{\hbar \Delta t} \int dx \rho(x, t) m^{AB} (\hbar \partial_A \phi - \dot{\xi}_A) (\hbar \partial_B \phi - \dot{\xi}_B) , \quad (17)$$

where $\Delta t = t' - t$. The divergence as $\Delta t \rightarrow 0$ is a consequence of the fact that as $\Delta t \rightarrow 0$ the distributions $P(x', t' | x, t)$ become infinitely narrow and therefore infinitely distinguishable. The problem is mildly annoying but not fatal because our interest is not in G itself but in how it changes as we vary $\Delta \xi^A$; it can be alleviated by choosing the arbitrary constant $C = \hbar \Delta t / 2$ so that

$$G = \frac{3N\hbar}{4\Delta t} + \frac{1}{2} \int dx \rho m^{AB} (\hbar \partial_A \phi - \dot{\xi}_A) (\hbar \partial_B \phi - \dot{\xi}_B) . \quad (18)$$

Using (13) to write ϕ in terms of Φ we find,

$$G = \frac{3N\hbar}{4\Delta t} - \frac{\hbar}{2} \dot{S} + \tilde{H}_0[\rho, \Phi] . \quad (19)$$

The first term is not interesting; it is a constant independent of $\dot{\xi}_A$. The second term is a bit more interesting because \dot{S} turns out to be the rate of entropy increase,

$$\dot{S} = \frac{dS}{dt} \quad \text{with} \quad S[\rho] = - \int dx \rho(x, t) \log \rho(x, t) . \quad (20)$$

Its contribution to the metric G is what one might expect: when $\dot{S} > 0$ the distributions are broader, more difficult to discriminate, and the information distance between them decreases. However \dot{S} is also independent of $\dot{\xi}_A$ and does not contribute to best matching. The explicit expression is

$$\dot{S} = - \int dx \rho \partial^A \left(\partial_A \Phi - \dot{\xi}_A \right) = - \int dx \rho \partial^A \partial_A \Phi , \quad (21)$$

where we used $\partial_A \dot{\xi}_B(t) = 0$. The term in (19) that is crucial for EBM is the last one,

$$\tilde{H}_0[\rho, \Phi] = \int dx \left[\frac{m^{AB}}{2} \rho \left(\partial_A \Phi - \dot{\xi}_A \right) \left(\partial_B \Phi - \dot{\xi}_B \right) + \frac{\hbar^2}{8} \frac{1}{\rho} \partial_A \rho \partial_B \rho \right] , \quad (22)$$

which can be recognized as the ensemble Hamiltonian (minus a potential energy term). Thus, considerations of information geometry have lead to a metric that includes the kinetic energy and the quantum potential. The significance of this finding will be more fully explored elsewhere.

Entropic Best Matching Entropic BM is achieved by minimizing the mismatch between $\rho(x', t'; x, t)$ and $\rho(x', t' + dt; x, t)$ as measured by the distance dT . This amounts to minimizing G with respect to trial shifts $\dot{\xi}^a$. The result of this variation is the constraint:

$$\frac{\partial G}{\partial \dot{\xi}^a} = \int dx \rho \sum_n \frac{\partial \Phi}{\partial x_n^a} - M \dot{\xi}_a^{\text{best}} = 0, \quad (23)$$

where we have used that $\dot{\xi}^A = \dot{\xi}^a$ is independent of x and of n and $M = \sum_n m_n$. Incidentally, (23) confirms our earlier intuition that $\Delta \xi^{\text{best}} \sim O(\Delta t)$.

The interpretation of $\dot{\xi}_a^{\text{best}}$ is straightforward. As discussed in [4] the integral

$$\tilde{P}_a = \int d^{3N}x \rho \sum_n \frac{\partial \Phi}{\partial x_n^a} = \int d^{3N}x \rho \frac{\partial \Phi}{\partial X^a} \quad (24)$$

is interpreted as the expectation of the total momentum, and X^a are the coordinates of the center of mass,

$$X^a = \frac{1}{M} \sum_n m_n x_n^a. \quad (25)$$

Equation (23) can now be read in two ways. If we are given a sequence of consecutive states $\{\rho, \Phi\}$ the shift $\dot{\xi}_a^{\text{best}}$ that achieves equilocality is given by the velocity of the center of mass,

$$M \dot{\xi}_a^{\text{best}} = \tilde{P}_a. \quad (26)$$

Alternatively, we can require that consecutive states be best matched, that is, no shift is needed: $\dot{\xi}_a^{\text{best}} = 0$. This means that successive states $\{\rho, \Phi\}$ are constrained to satisfy $\tilde{P}_a = 0$. In this second reading (23) is a constraint on the allowed states.

4 Relational Quantum Dynamics

The dynamics described by eq.(12) for any externally prescribed drift potential ϕ is a diffusion; it is an unusual diffusion in that it is relational, but it is not yet a *quantum* theory. The way to a quantum mechanics is to require that the diffusion be non-dissipative which is what leads to a Hamiltonian dynamics.

The idea is to recognize that, just as ρ evolves in response to ϕ , we must also allow ρ to react back so that ϕ evolves in response to ρ . This promotes φ , or equivalently Φ , to a dynamical variable. The precise recipe for this coupled evolution is to require that there be some conserved quantity $\tilde{H}[\rho, \Phi]$ such that changes in ρ induce changes in Φ in a way that $\tilde{H}[\rho, \Phi]$ remains constant. The quantity $\tilde{H}[\rho, \Phi]$ is chosen so that the evolution of ρ is described by eq.(12). The result is the coupled evolution of ρ and Φ given by Hamilton's equations [3],

$$\partial_t \rho = \frac{\delta \tilde{H}[\rho, \Phi]}{\delta \Phi} \quad \text{and} \quad \partial_t \Phi = -\frac{\delta \tilde{H}[\rho, \Phi]}{\delta \rho}. \quad (27)$$

The Hamiltonian $\tilde{H}[\rho, \Phi]$ that reproduces (12) is of the form,

$$\tilde{H}[\rho, \Phi] = \int dx \frac{\rho}{2} m^{AB} \left(\partial_A \Phi - \dot{\xi}_A \right) \left(\partial_B \Phi - \dot{\xi}_B \right) + F[\rho], \quad (28)$$

where the functional $F[\rho]$ is an integration constant. The second Hamilton equation is recognized as a Hamilton–Jacobi (HJ) equation,

$$\begin{aligned} -\partial_t \Phi &= \frac{1}{2} m^{AB} \left(\partial_A \Phi - \dot{\xi}_A \right) \left(\partial_B \Phi - \dot{\xi}_B \right) + \frac{\delta F}{\delta \rho} \\ &= \sum_n \frac{\delta^{ab}}{2m_n} \left(\frac{\partial \Phi}{\partial x_n^a} - m_n \dot{\xi}_a \right) \left(\frac{\partial \Phi}{\partial x_n^b} - m_n \dot{\xi}_b \right) + \frac{\delta F}{\delta \rho} . \end{aligned} \quad (29)$$

$\partial \Phi / \partial x_n^a$ is interpreted as the momentum p_a of the n^{th} particle. In fact, p_a is the “coordinate” momentum; the intrinsic momentum that reflects true change is the momentum corrected by the shift velocity, $p_a - m_n \dot{\xi}_a$.⁶

The advantage of EBM over classical BM is that the framework introduced here is quantized in a straightforward fashion. In ED quantum theory results from a particular choice of the functional $F[\rho]$. What is especially interesting is that this special choice is suggested by the information metric in eq.(19) and specifically by the term \tilde{H}_0 , eq.(22). By choosing $F[\rho]$ so that the Hamiltonian coincides with eq.(22) we indeed obtain a quantum theory, a theory of N free particles. More interesting quantum theories are obtained by the inclusion of some scalar potential $U(x)$. This results in the quantum Hamiltonian,

$$\tilde{H}[\rho, \Phi] = \tilde{H}_0[\rho, \Phi] + \int dx \rho U . \quad (30)$$

The fact that this is now a *quantum* theory can be made explicit by combining ρ and Φ into a complex wave function $\Psi = \rho^{1/2} \exp(i\Phi/\hbar)$. Then Hamilton’s equations (27) can be combined into a single linear Schrödinger equation,

$$i\hbar \partial_t \Psi = \frac{1}{2} m^{AB} \left(i\hbar \partial_A - \dot{\xi}_A \right) \left(i\hbar \partial_B - \dot{\xi}_B \right) \Psi + U(x) \Psi . \quad (31)$$

If we require that consecutive states be best matched, that is we impose $\dot{\xi}_a = 0$, then the states $\{\rho, \Phi\}$ are constrained to satisfy $\hat{P}_a = 0$. In terms of Ψ this constraint takes the form

$$\int dx \Psi^*(x, t) \sum_n \hat{p}_{na} \Psi(x, t) = \langle \hat{P}_a \rangle = 0, \quad (32)$$

where $\hat{p}_{na} = i\hbar \partial / \partial x_n^a$ is the quantum momentum operator for particle n and $\hat{P}_a = \sum_n \hat{p}_{na}$. Thus, EBM imposes a constraint that restricts solutions to the subspace of the full Hilbert space where the expected total momentum vanishes.

As a side remark, it is interesting that EBM imposes *weak* constraints, or rather, expected value constraints. In the standard approach to quantizing theories with constraints, questions arise as to whether constraints should be imposed on operators, on states, or on expectation values. EBM provides an answer — quantum constraints are to be imposed on expectation values.

⁶Something analogous occurs in electromagnetism. The canonical momentum p^a is split into a kinetic momentum π^a and the momentum contained in the electromagnetic field, eA^a/c . Therefore $p^a = \pi^a + eA^a/c$. The momentum mv^a that contributes to the kinetic energy $mv^2/2$ is $mv^a = \pi^a = p^a - eA^a/c$.

To complete the relational ED we require yet another condition, a consistency condition. For the dynamics to remain relational we must require that the potential $U(x)$ be such that the constraint $\langle \hat{P}_a \rangle = 0$ be preserved by the dynamical evolution. This implies that $\langle \hat{P}_a \rangle$ must be a constant of the motion. This condition is satisfied if the potential $U(x)$ depends only on the relative particle positions, $U(x) = U(\{\vec{x}_i - \vec{x}_j\})$. The non-conservation of momentum would indicate the existence of a preferred absolute frame of reference.

5 Discussion and Conclusion

We have formulated a relational quantum dynamics by implementing an entropic form of best matching. The framework has been developed for the simple case of global translations in particle dynamics but it can be easily generalized to global rotations in which case the imposed constraint is that the total expected angular momentum vanish. That these results are statistical analogues of the classical BM is both encouraging and deceiving. On one hand, one expects EBM to reduce to classical BM in some limit, but the two frameworks are very different animals. Classical BM compares ontic particle configurations while EBM compares epistemic probability distributions. The difference is highlighted in the case of a single particle. Classically we cannot even ask about the relational motion of a single particle. Relative to what would that motion happen? The situation in EBM is appreciably different, there is no problem either in principle or in practice for even a single particle because it is probability distributions that are being best-matched and not the particle configurations. Another important difference is that classical BM relies on the classical action as a best-matching criterion, EBM relies on a purely inferential criterion provided by information geometry. In fact, while classical BM requires previous knowledge of the dynamics, in this work we have uncovered hints that EBM will help us find the Hamiltonian that induces the dynamics.

To conclude, the theory developed here is not yet fully relational. This is only a first step towards a relational framework that applies to the local gauge symmetries including diffeomorphisms and local dilatations that are central to all fundamental theories such as electromagnetism, Yang-Mills theories, and gravity.

Acknowledgements

We would like to thank M. Abedi, D. Bartolomeo, C. Cafaro, N. Caticha, S. DiFranzo, A. Giffin, D.T. Johnson, K. Knuth, S. Nawaz, M. Reginatto, C. Rodríguez, and K. Vanslette for many discussions on entropy, inference and quantum mechanics.

References

- [1] A. Caticha, “Entropic dynamics, time, and quantum theory”, J. Phys. A: Math. Theor. **2011**, 44, 225303; arXiv.org:1005.2357.
- [2] A. Caticha, “Entropic Dynamics: an inference approach to quantum theory, time, and measurement”, J. Phys.: Conf. Ser. **2014**, 504, 012009; arXiv.org:1403.3822.
- [3] A. Caticha, D. Bartolomeo, M. Reginatto, “Entropic Dynamics: from entropy and information geometry to Hamiltonians and quantum mechanics”, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, ed. by A. Mohammad-Djafari, and F. Barbaresco, AIP Conf. Proc. **1641**, 155 (2015); arXiv.org:1412.5629.
- [4] A. Caticha, “Entropic Dynamics”, Entropy **17**, 6110 (2015); arXiv.org:1509.03222.
- [5] For a pedagogical review see A. Caticha, *Entropic Inference and the Foundations of Physics* (monograph commissioned by the 11th Brazilian Meeting on Bayesian Statistics – EBEB-2012); online at <http://www.albany.edu/physics/ACaticha-EIFP-book.pdf>.
- [6] S. Nawaz, M. Abedi, and A. Caticha, “Entropic Dynamics on Curved Spaces”, in these Proceedings.
- [7] A. Caticha, “Geometry from Information Geometry”, in these Proceedings; arXiv:1512.09076.
- [8] J. B. Barbour and B. Bertotti, “Mach’s Principle and the Structure of Dynamical Theories”, Proc. Roy. Soc. A **382**, 295 (1982).
- [9] Many excellent papers by J. B. Barbour and his collaborators are found online at platonica.com. For a pedagogical review with extensive references see F. Mercati, “A Shape Dynamics Tutorial”, arXiv:1409.0105 [gr-qc].
- [10] D. Bartolomeo and A. Caticha, “Entropic Dynamics: The Schrödinger equation and its Bohmian limit”, in these Proceedings; arXiv:1512.09084.
- [11] S. B. Gryb, “Implementing Mach’s Principle Using Gauge Theory”, Phys. Rev. D **80**, 024018 (2009); arXiv:0901.2362 [gr-qc].
- [12] S. Amari, *Differential-Geometrical Methods in Statistics* (Springer-Verlag, 1985).